TOPOLOGY OF THE MEMBRANE-BOUND ALKANE HYDROXYLASE OF *PSEUDOMONAS OLEOVORANS*

CHAPTER 4

Journal of Biological Chemistry **267**: 9194-9201 (1992)

Jan B. van Beilen, Dirk Penninga and Bernard Witholt

TOPOLOGY OF THE MEMBRANE-BOUND ALKANE HYDROXYLASE OF *PSEUDOMONAS OLEOVORANS*

SUMMARY

The *Pseudomonas oleovorans* alkane hydroxylase is an integral cytoplasmic membrane protein which is expressed and active both in *Escherichia coli* and *Pseudomonas oleovorans*. Its primary sequence contains eight hydrophobic stretches which could span the membrane as α-helices. The topology of alkane hydroxylase was studied in *E. coli* using protein fusions linking different amino-terminal fragments of the alkane hydroxylase (AlkB) to alkaline phosphatase (PhoA) and to ß-galactosidase (LacZ).

Four AlkB-PhoA fusions were constructed using transposon Tn*phoA*. Site-directed mutagenesis was used to create *Pst*I sites at 12 positions in AlkB. These sites were used to create AlkB-PhoA and AlkB-LacZ fusions. With respect to alkaline phosphatase and ß-galactosidase activity each set of AlkB-PhoA and AlkB-LacZ fusions revealed the expected complementary activities. At three positions, PhoA fusions were highly active while the corresponding LacZ fusions were the least active. At all other positions, the PhoA fusions were almost completely inactive while the corresponding LacZ fusions were highly active.

These data predict a model for alkane hydroxylase containing six transmembrane segments. In this model the amino terminus, two hydrophilic loops, and a large carboxy terminal domain are located in the cytoplasm. Only three very short loops near aminoacid positions 52, 112 and 251 are exposed to the periplasm.

INTRODUCTION

Aliphatic hydrocarbons can be degraded by a wide range of microorganisms. One of these is the Gram-negative *Pseudomonas oleovorans* which can utilize medium-chain length alkanes by virtue of two operons, located on the OCT-plasmid (Chakrabarty *et al.*, 1973). The *alkBFGHJKL* operon encodes two components of the alkane hydroxylase complex (Benson *et al.*, 1977; Peterson *et al.*, 1966) and other enzymes in the alkane oxidation pathway (Fennewald *et al.*, 1979; Eggink *et al.*, 1987a). The *alkST* operon encodes a regulator (AlkS) of *alkBFGHJKL* expression (Fennewald and Shapiro, 1977; Eggink *et al.*, 1988), and the third component of the alkane hydroxylase complex.

The components of the alkane hydroxylase complex are rubredoxin reductase (AlkT) (Eggink *et al.*, 1990; Ueda *et al.*, 1972), which transfers electrons from NADH to rubredoxin; rubredoxin (AlkG) (Kok *et al.*, 1989b; Peterson *et al.*, 1966), which is an iron-sulfur electron-transfer protein; and AlkB (Kok *et al.*, 1989a), a monooxygenase which catalyzes the oxidation of alkanes with one oxygen atom derived from molecular oxygen. The second oxygen atom is reduced with the electrons transferred from NADH via rubredoxin reductase and rubredoxin.

Immunolocalization and cell fractionation experiments have shown that AlkB is associated with the cytoplasmic membrane (Benson *et al.*, 1979; Lageveen, 1986), in accordance with previous experiments which showed that alkane hydroxylase requires phospholipids for catalytic activity (Ruettinger *et al.*, 1974).

We have recently cloned and sequenced *alkB* (Kok *et al.*, 1989a). The primary structure of AlkB contains eight hydrophobic stretches of

about 20 amino acid residues, each of which could in principle span the cytoplasmic membrane, suggesting that AlkB is an integral membrane protein.

In this paper we analyze the primary sequence of AlkB to develop a model for the transmembrane topology of this enzyme. To test and refine the preliminary model we have made a series of alkaline phosphatase (PhoA) and ß-galactosidase (LacZ) fusions to AlkB at specific sites, to determine whether individual sites are located on or near the cytoplasmic or periplasmic membrane face. This approach is based on the finding that PhoA lacking its own export signal is active when fused to periplasmic domains but exhibits little or no activity when fused to cytoplasmic domains of membrane proteins (Michaelis *et al.*, 1983). LacZ fusions provide complementary information; this enzyme is active when fused to cytoplasmic domains but shows decreased activity when it is fused to periplasmic domains of membrane proteins (Bassford *et al.* 1979; Froshauer *et al.* 1988).

EXPERIMENTAL PROCEDURES

Materials- The reagents for bacterial cell growth were purchased from Difco. The chromogenic substrates 5-bromo-4-chloro-3-indolyl phosphate (XP), p-nitrophenylphosphate (Sigma 104), o-nitrophenyl-ß-d-galactose (ONPG) and goat antirabbit IgG coupled to alkaline phosphatase were purchased from Sigma. Restriction enzymes, DNA polymerase I (Klenow fragment), T4 DNA ligase, polynucleotide kinase and dideoxynucleotides were from Boehringer Mannheim, radionucleotides were from Amersham Corp, and deoxynucleotides used in sequencing and in site-directed mutagenesis experiments were purchased from Pharmacia LKB Biotechnology Inc. Oligonucleotides were synthesized by Eurosequence.

Protein and DNA sequence analysis- Protein and DNA sequence analysis was carried out using programmes collected in the PC/GENE software package (IntelliGenetics, Geel, Belgium), developed by A. Bairoch.

Bacterial Strains, Plasmids, Phages and Growth Media- E. coli K12 strains, phages and plasmids used in this study are listed in Table I. Plasmid pBG201 is a pUC19 derivative in which a 2282 bp SmaI-HindIII fragment from pGEc48 containing the *alkB* gene including the *alk* promoter is cloned. Plasmid pBG203 is a pGEM-7Zf(+) derivative in which the same fragment is cloned as an EcoRI-HindIII fragment from pBG201. Plasmid pGEc74 encodes AlkS which is necessary for transcription of the *alkB* gene. Bacteriophage λTn*phoA* was used as a source of Tn*phoA*. λTn*phoA* was propagated in *E. coli* K-12 LE372. Plasmid pGEc47ΔB is an *alkB* deletion derivative of pGEc47 (Eggink *et al.*, 1987b) which contains all of the *alk* system. A 528 basepair *Bam*HI-*Bam*HI fragment was first deleted from the *alkB* gene in pGEc48. The deletion was then transferred to pGEc47 by marker rescue using the *E. coli polA-* strain SF800.

LB (Maniatis) and E-2 medium (Eggink *et al.*, 1987b), supplemented with carbon sources or antibiotics were used throughout. To induce alkaline phosphate in *E. coli* JM101 cells were grown in a Tris-buffered medium containing 0.5 mM inorganic phosphate (Garen and Levinthal, 1960). Antibiotics were used at the following final concentrations: ampicillin, 100 μg/ml; kanamycin, 50 μg/ml; tetracycline, 12.5 μg/ml. All cultures were grown aerobically at 30°C or 37°C.

*Isolation and characterization of Tn*phoA *insertions-* Alkaline phosphatase gene fusions were obtained by infecting cells (CC118 carrying plasmid pBG201) with bacteriophage λTn*phoA*. Briefly, cells were grown at 37°C to early stationary phase in LB containing 10 mM magnesium chloride and 100 ug/ml ampicillin. The phage λTn*phoA* was added to the culture at a multiplicity of one and the incubation was continued for 15 min at 30°C. The culture was diluted 10-fold with LB medium and grown

Table I
List of strains, bacteriophages, and plasmids

E. coli Strain	Relevant genotype or description	Source or reference
CC118	*phoAΔ20*	Manoil and Beckwith, 1985
LE392	*supF, supE*	Silhavy *et al.*, 1984
BW313	*dut, ung*	Kunkel *et al.*, 1987
JM101	F⁻ *(proAB, lacI*^{q}, *lacZM15), thi-1, Δ(lac-proAB)*	Yanish-Perron *et al.*, 1985
DH-1	*recA1*	Maniatis *et al.*, 1982
X7026	*Δ(gpt-lac)5*	Kahn, 1968
GEc137	DH-1 *fadR*	Eggink *et al.*, 1987b
SF800	W3110 *polA*	Laboratory collection
Phage or plasmid	**Replicon, markers**	**Source or reference**
λTn*phoA*	*cI857*, $P_{am}3$ with Tn*phoA* in/near *rex*	Gutierrez *et al.*, 1987
M13KO7	M13 with Km resisitance gene and P15A origin inserted in the intergenic region	Vieira and Messing, 1987
pUC18	Ap, ColE1	Yanish-Perron *et al.*, 1985
pGEM-7Zf(+)	Ap, ColE1, M13 intergenic region	Promega
pSKS105	Ap, *lacZYA*, ColEI	Casabadan *et al.*, 1983
pMC1871	Tc, *lacZ*, ColE1	Shapira *et al.*, 1983
pCH2	Tc, *phoA*, ColE1	Hoffman and Wright, 1985
pCH40	Tc, *phoA*, ColE1	Hoffman and Wright, 1985
pGEc47	Tc, *alkBFGHJKL/alkST*, RK2	Eggink *et al.*, 1987b
pGEc47ΔB	pGEc47, 528 bp *Bam*HI del. in *alkB*	this study
pGEc48	*alkBFGH* in pBR322, ColE1, Ap	Eggink *et al.*, 1987a
pGEc74	*alkST* in pLAFR1, RK2, Tc	Eggink *et al.*, 1988
pBG201	*alkBF* in pUC18	this study
pBG203	*alkBF* in pGEM-7Zf(+)	this study
pBG203P(X)	pBG203, *Pst*I sites P(X) X=1-12	this study
pAP(X)	pBG203P(X), with *Pst*I-*Hin*dIII fragment replaced by *phoA*	this study
pAP1N	pAP1, with introduced *Nsi*I site	this study
pAP(X)N	pBG203P(X), with *Pst*I-*Hin*dIII fragment replaced by *Nsi*I-*Hin*dIII fragment from pAP1N X=3,4,6,7,8,9	this study
pLacZ(X)	pBG203P(X), with *lacZ* fragment inserted in *Pst*I sites X=1-12	this study

Ap, ampicillin resistance; Km, kanamycin resistance; Tc, tetracyclin resistance

overnight at 30°C. Independently infected cultures were plated on TYE-plates containing 300 μg/ml kanamycin, 200 μg/ml ampicillin and 40 μg/ml XP to select Tn*phoA* transpositions onto the multicopy plasmid pBG201. After two days all colonies were pooled, and plasmid DNA was prepared. To selectively recover those plasmids into which transposition had occurred the DNA was used to transform competent CC118 cells carrying plasmid pGEc74, which is necessary for transcription of the *alkB* gene, as it contains the regulatory gene *alkS*. Transformants were selected at 37°C on LB plates containing 50 μg/ml kanamycin and 40 μg/ml XP. All blue colonies were transferred to TYE containing 50 μg/ml kanamycin, 100 μg/ml ampicillin and 40 μg/ml XP, and plasmid DNA from ampicillin resistant transformants was isolated. This procedure was necessary because selection of transformants directly on LB plates containing both ampicillin and kanamycin resulted in many dark blue colonies which contained both the original plasmid with an intact ß-lactamase gene, responsible for ampicillin resistance, and plasmids in which the transposon had inserted in the ß-lactamase gene, responsible for the kanamycin resistance and PhoA+ phenotype.

*Eco*RI digestion of selected plasmids was used to determine the positions of Tn*phoA* insertions. Plasmids were stabilized by deleting most of the Tn*5* sequences, including the transposase and the *npt* gene, using *Xho*I. The exact fusion point was determined by the Sanger dideoxynucleotide chain termination method on double stranded DNA using a 16-mer-Tn*phoA* primer corresponding to a sequence beginning 80 nucleotides into Tn*phoA*. The resulting gene fusions and their characteristics are shown in Table II.

Introduction of PstI sites in the alkB sequence- To construct *phoA* and *lacZ* fusions to selected sites in *alkB*, we introduced *Pst*I sites at the desired locations (Chapter 6, Table II). BW313 was used to isolate single-stranded template DNA containing uridine (Kunkel *et al.*, 1987) from phagemid clones (McClary *et al.*, 1989) for site-directed mutagenesis. On average more than 50% of plasmids resulting from these mutagenesis experiments contained the desired mutation (resulting in plasmids pBG203P1-P12).

In most cases the introduction of *Pst*I sites in the *alkB* gene led to changes in the amino acid sequence. To test whether the introduction of the *Pst*I sites had inactivated AlkB, *E. coli* GEc137 carrying plasmid pGEc47ΔB (all *alk* genes necessary for growth on octane, except *alkB*) was transformed with pBG203P1-12. The resulting transformants were streaked on E2 minimal media plates supplemented with thiamine. The Petri dishes were incubated at 32°C in a sealed container under n-octane vapor for 7 days to test growth on n-octane.

Construction of alkB-phoA and alkB-lacZ gene fusions- The truncated *phoA* gene was introduced in the plasmids containing the *Pst*I sites as follows: A *Pst*I-*Xho*I fragment containing the *phoA* gene was isolated from plasmid pCH2 for *Pst*I sites 1 to 9, or pCH40 for *Pst*I sites 10 to 12, depending on the position of the *Pst*I site relative to the AlkB reading frame, and the *Xho*I sticky end was filled in with Klenow DNA polymerase. This fragment was cloned between the *Pst*I and *Eco*RV sites of the mutagenized plasmids (resulting in plasmids pAP1-12, encoding fusion proteins AP1-12).

To obtain AlkB-LacZ fusions a *Pst*I fragment from plasmid pMC1871 was inserted in the created *Pst*I sites of pBG203P1-9. For pBG203P10-12 a plasmid carrying a *lacZ* gene with the *Pst*I site in the appropriate frame was constructed by cloning an *Eco*RI-*Sst*I fragment from pSKS105 (first half of *lacZ*) and a *Sst*I-*Pst*I fragment from pMC1871 (last half of *lacZ*) between the *Eco*RI and *Pst*I sites of pUC18. The AlkB-LacZ fusion plasmids were named pLacZ1-12, encoding fusion proteins BG1-12. The orientation of the inserts was determined by restriction enzyme mapping.

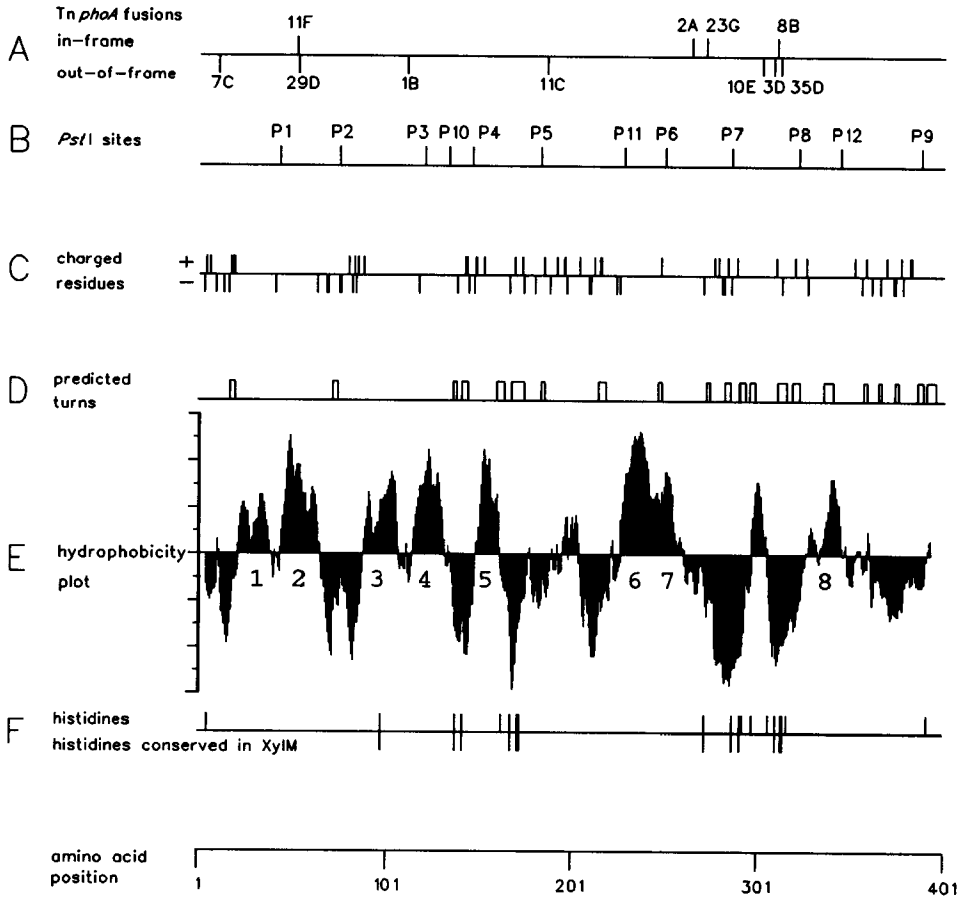

Fig. 1. **Location of Tn*phoA* transposition, introduced *Pst*I sites, charged residues and histidines, and the hydrophobicity plot of AlkB.** The scale at the bottom line indicates the residue numbers in the AlkB sequence. **A.** Tn*phoA* transpositions. In-frame fusions are shown above the axis, out-of-frame fusions below the axis. **B.** *Pst*I sites introduced into the *alkB* sequence by site-directed mutagenesis. **C.** Charged residues: vertical lines above the axis indicate lysines and arginines, vertical lines below the axis indicate aspartic acid and glutamic acid residues. **D.** Segments with high turn propensity identified according to Paul and Rosenbusch (1985) **E.** Hydropathy profile of the AlkB sequence. The profile was constructed using the Kyte-Doolittle algorithm (1982) with a sliding window of seven amino acids. Values above the zero axis correspond to hydrophobic segments. The numbers 1-8 refer to 8 possible membrane-spanning domains. **F.** Histidines: vertical lines above the axis indicate all histidines in AlkB, vertical lines below the axis indicate histidines conserved in XylM.

Introduction of hydrophobic stretch 1 between the AlkB and PhoA domains of fusion proteins AP3 to AP9- An *Nsi*I site was introduced at amino acid position 10 of plasmid pAP1 (resulting in plasmid pAP1Nsi) to study the properties of hydrophobic segment 1 in different contexts. The *Nsi*I-*Hin*dIII fragment from pAP1Nsi, encoding PhoA attached to AlkB amino acids 10-44, was cloned between the *Pst*I and *Hin*dIII sites of the plasmids containing the P3, P4, P6, P7, P8, and P9 *Pst*I sites in the *alkB* gene (resulting in plasmids pAP3-9N, encoding fusion proteins AP3-9N). AP(X)N differs from AP(X) in that hydrophobic stretch 1 is introduced in the *Pst*I site between the AlkB en PhoA moiety of the fusion protein.

Alkaline phosphatase and β-galactosidase assays- Alkaline phosphatase and ß-galactosidase activity were determined essentially as described before (Brickman and Beckwith, 1975; Miller, 1972) using p-nitrophenylphosphate and o-nitrophenyl-ß-d-galactose, respectively. Enzymatic activity was expressed in Miller units (Miller, 1972). For both assays, L-medium cultures were inoculated directly from fresh plates to optical density at 600 nm of about 0.1, grown further for 3 h at 30°C to bring them into exponential-phase growth, induced with 0.1% DCPK, and grown for an additional 3 h. The optical density at 600 nm was determined (generally between 1.0 and 2.0). The 2 ml cultures were harvested by centrifugation, and then resuspended in 1 ml 1 M Tris-HCl pH 8.0 (PhoA) or Z-buffer (LacZ). Cells were permeabilized with 50 μl 0.1% SDS and 50 μl $CHCl_3$. Reactions were started by the addition of 0.1 ml 0.4% p-nitrophenylphosphate (PhoA) or o-nitrophenyl-ß-d-galactoside (LacZ), and stopped by the addition of 0.1 ml 1.0 M K_2HPO_4 (PhoA) or 0.5 ml 1 M sodium carbonate (LacZ), and placing the tubes on ice. After centrifugation in a microfuge, the OD_{405} of the supernatant was measured. Enzymatic activity was expressed in Miller units (Miller, 1972).

Western blotting and Immunostaining- SDS-PAA gel electrophoresis was done according to Laemmli (1972). Cells were grown as described above, harvested by centrifugation and resuspended in H_2O. Samples were mixed with 4 volumes of loading buffer, boiled for 5 min (Silhavy *et al.*, 1984) and subjected to gel electrophoresis. Proteins were transferred from the gel to nitrocellulose using an LKB semi-dry blotting apparatus. Immunostaining was carried out using rabbit anti-PhoA IgG (Jan Tommassen), or rabbit anti-LacZ serum (Greetje Vos-Scheperkeuter), followed by PhoA-coupled goat antirabbit IgG antibodies (Sigma). Sizes were estimated by negative staining of the marker proteins with india ink (Harlow and Lane, 1988).

RESULTS

Primary structure analysis of the AlkB sequence- The amino acid sequence of AlkB was analyzed using several algorithms (programs collected in the PC/GENE software package) which detect membrane-spanning or membrane associated sequences. The procedure of Kyte and Doolittle (1982) for the distribution of hydrophobic amino acids sequences results in the hydropathy profile shown in Fig. 1E. AlkB has 8 sequences which could span the cytoplasmic membrane. The hydrophobicity patterns between residues 20 and 62, 89 and 134, and 228 and 271 strongly suggest that each of these regions spans the membrane twice. The regions between residues 150 and 166, and between residues 329 and 352 are quite hydrophobic as well.

Each of these stretches is classified as a potential transmembrane or membrane associated sequence by at least one of the algorithms that are based on hydrophobic properties. However, these results are not unequivocal, because only 3 stretches (2, 4 and 7) are classified as membrane spanning by all of the algorithms used. We have therefore turned to protein fusion

techniques to determine which of the hydrophobic stretches actually span the cytoplasmic membrane.

Gene Fusions Constructed by TnphoA Transposition- Alkaline phosphatase gene fusions were obtained by infecting cells (CC118 carrying plasmid pBG201) with bacteriophage λTn*phoA* (Gutierrez *et al.*, 1987) and selecting for insertion of the transposon in pBG201 as described before (Manoil and Beckwith, 1985). Selection for plasmids encoding active AlkB-PhoA fusions was carried out in CC118 carrying plasmid pGEc74, which contains the *alkS* gene, necessary for transcription of the *alkB* gene and *alkB-phoA* gene fusions. Colonies producing active AlkB-PhoA fusion proteins were identified by the formation of pale blue to dark blue colonies on media containing the indicator dye 5-chloro-4-bromo-3-indolyl phosphate.

The restriction patterns of 186 independent *alkB-phoA* gene fusions were determined. Most insertions (80%) that exhibited a PhoA$^+$ phenotype had inserted in the wrong orientation or outside *alkB*, presumably in the ß-lactamase gene. 36 Fusions were located in *alkB* in the correct orientation. In these cases the exact fusion point was determined by DNA sequencing. The fusions occurred at only 11 unique positions. Of these, only four were in-frame-fusions. These were located in hydrophobic stretch 2, near the middle of stretch 7, near the end of stretch 7, and in between stretch 7 and 8, respectively (Fig. 1A). The first fusion was reasonably active, while the other fusions showed low alkaline phosphatase activities (Table II).

Of the out-of-frame fusions, six were in the +1 frame, and one was in the +2 frame relative to the AlkB reading-frame (Fig. 1A). Most of these were inactive, and resulted in pale blue colonies. However, three +1 out-of-frame fusions showed variable activities (Table II) and yielded dark blue colonies. A Western blot of the two most active out-of-frame fusions produced an immunoreactive band with a molecular

Table II

*List of alkB-phoA gene fusions constructed by random transposition of Tn*phoA

Allele Code	Amino Acid Position	Nucleotide Sequence around Fusion Joints[a]	PhoA$^+$ activity[b]
11F	52 in frame	CTG.GTA.T *TG.CTC.GTA	135
2A	265 in frame	TGG.CAG.C *TG.ACC.AGT	40
23G	273 in frame	ATT.GAA.C *AT.TAC.GGC	25
8B	311 in frame	CTT.CAG.C *GG.CAC.TCG	12
7C	10 o.f. +1	CTG.GAT.TC *C.GCT.CCA	3
29D	53 o.f. +1	GTA.TTG.CT *C.GTA.TGG	13 - 100
1B	111 o.f. +1	GTC.GGA.AC *T.CAG.CCA	3 - 40
10E	303 o.f. +1	TCT.AAT.CT *A.GTG.CTG	35 - 260
3D	309 o.f. +1	TTC.CAC.CT *T.CAG.CGG	3
35D	313 o.f. +1	CGG.CAC.TC *G.GAT.CAC	3
11C	187 o.f. +2	CGG.ATG.GGA *.GAA.AGC	4

[a]Exact amino acid position and frame of the fusions were determined by Sanger DNA sequencing. The * indicates the fusion joint in the *alkB* sequence. [b]Alkaline phosphatase activity of *E. coli* CC118 harbouring *alkB-phoA* fusion plasmids was measured in permeabilized cells with p-nitrophenyl phosphate as substrate (Miller units). 'o.f.'is out-of-frame.

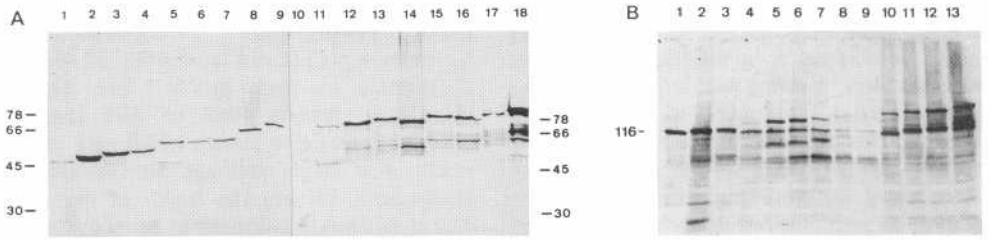

Fig. 2. **AlkB-PhoA and AlkB-LacZ fusion proteins as analyzed by Western blotting.** Panel A. AlkB-PhoA protein fusions were separated by electrophoresis on 10% SDS-PAGE, transferred to nitrocellulose and immunostained using anti-PhoA IgG. Total cell extracts were prepared as described in Methods and Materials. *E. coli* JM101 was induced for PhoA using low-phosphate medium (*lanes 1 and 10*). Strain CC118 carrying the AlkB-PhoA fusions was grown in LB-medium and induced with DCPK (*lanes 2-9*: fusion AP1, 11F, AP2, AP3, AP10, AP4, AP5, AP11, AP6, *lanes 11-18*: 2A, 23G, AP7, 8B, AP8, AP12, AP9). Panel B. AlkB-LacZ protein fusions were separated on 7.5% SDS-PAGE, transferred to nitrocellulose, and immunostained using anti-LacZ serum (*lane 1*: DH1 induced for LacZ, *lanes 2-13*: fusions BG1, BG2, BG3, BG10, BG4, BG5, BG11, BG6, BG7, BG8, BG12, BG9).

weight slightly larger than that of mature alkaline phosphatase (data not shown). Upon investigation of possible open-reading-frames in the +1 frame of the *alkB* gene, we found that both fusions were preceded by a reasonable Shine-Dalgarno box, an ATG startcodon, and a peptide resembling a signal sequence at their fusion point. Therefore, these out-of-frame fusions apparently led to new translation starts within *alkB* resulting in the synthesis and export of a near-native alkaline phosphatase. Consequently, these fusions had no relevance to the analysis of AlkB topology and were not considered further.

Since only 4 out of a total of 186 independent blue colonies which we analyzed represented unique in-frame *alkB-phoA* fusions, we turned to site-directed mutagenesis to create additional and specific *alkB-phoA* fusions, including PhoA fusions to cytoplasmic domains of AlkB. Because these fusions are inactive and cannot be distinguished from out-of-frame fusions or fusions outside the *alkB* gene they are very difficult to isolate using the Tn*phoA* approach.

Construction of alkB-phoA *and* alkB-lacZ *gene fusions by site-directed mutagenesis*- To obtain specific fusions, we introduced *Pst*I sites (P1-P12) at positions corresponding to all putative extra-membranous domains (Fig. 1B), as described in Chapter 6. This allowed us to construct both PhoA and LacZ fusions at these positions, designated AP1-12 and BG1-12 respectively.

For the three putative dimeric transmembrane sequences we introduced *Pst*I sites P1, P3 and P6 near the middle of the hydrophobic sequence, two or three amino acids distal of the charged residue that was present in the middle of all three sequences. Most of the other *Pst*I sites were placed near the middle of the hydrophilic segments, since Boyd and Beckwith (1989) have shown that fusions at the beginning of cytoplasmic loops can show 7-20 fold more alkaline phosphatase activity than those at the ends of cytoplasmic loops. PhoA fusions were constructed by replacing the *Pst*I-*Eco*RV fragment encoding the carboxy terminal part of AlkB by a *Pst*I-*Xho*I fragment from pCH2 or pCH40 (depending of the position of the *Pst*I site relative to the AlkB reading frame) encoding PhoA without its signal sequence. LacZ fusions were constructed by inserting a *Pst*I fragment encoding LacZ from pMC1871 or a vector derived from pSKS105 into the created *Pst*I sites.

The introduction of the *Pst*I sites resulted in the loss of alkane hydroxylase activity only for sites P5, P7 and P8 (Chapter 6). For all of the fusions

created at each of the *PstI* sites at least one of the two fused enzymatic probes was active and could be detected as a full-length product in Western blots, showing that during the mutagenesis procedure no frame-shift mutations had occurred outside the region that was sequenced to test the mutations. Several protein fusions were tested for remaining alkane hydroxylase activity as well. Interestingly, fusion BG9 was able to complement an *alkB* deletion, unlike fusion AP9, and all fusions at other positions (data not shown).

Translation products of AlkB-PhoA and AlkB-LacZ fusions- With the exception of PhoA-fusions AP1 and 11F, hybrid AlkB-PhoA proteins could not be detected in whole cell extracts on Coomassie-blue stained SDS-polyacrylamide gels (data not shown). Therefore, we carried out Western blots using anti-PhoA antibodies to establish the apparent molecular weights of the hybrid proteins (Fig. 2A). The sizes of the fusion proteins were as predicted from the DNA sequence. Tn*phoA* fusions are 2.75 kDa (24 amino acids) larger than corresponding site-directed PhoA fusions because of the length of the Tn*5* spacer.

Alkaline phosphatase fusions 11F, AP3 and AP6 show a degradation product the size of mature alkaline phosphatase. This is probably due to processing by the *E. coli* signal peptidase, indicating a periplasmic localization of alkaline phosphatase in these fusions. The alkaline phosphatase fusions distal of AP6 show a different degradation pattern which may be the result of a single cleavage site in alkaline phosphatase, because the length of the main degradation products appears to be related to the length of the AlkB portion of the fusion protein.

Most of the AlkB-LacZ fusion proteins were visible on a Coomassie stained gel, with the exception of BG3, BG11 and BG6 (data not shown). Of all AlkB-LacZ fusion proteins BG1 is the most abundantly produced of the whole set, like the corresponding PhoA fusion. On a Western blot (Fig. 2B) only BG6 is not visible as a full-length fusion protein. BG3 and BG11 produce faint bands. Apparently the fusion proteins can be cleaved at several positions within the LacZ domain, resulting in products of the same length for all fusion proteins (LacZ fragments), while the length of other fragments is correlated to the size of the AlkB portion of the fusion protein (AlkB-LacZ' fragments).

Enzyme activity of AlkB-PhoA and AlkB-LacZ fusions- The alkaline phosphatase and ß-galactosidase activities of each fusion protein are

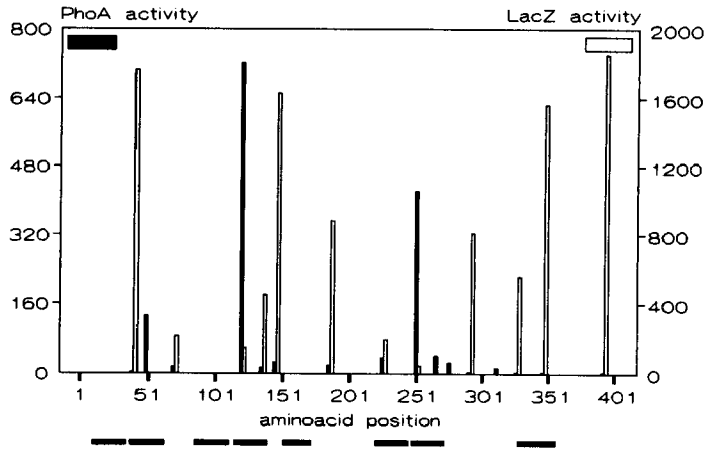

Fig. 3. **Alkaline phosphatase and ß-galactosidase activity of cells expressing fusion proteins.** This bar graph was constructed from data presented in Table III. The enzyme activities were determined as described under "Experimental Procedures" and then plotted against the amino acid position in AlkB of the gene fusion. The horizontal black bars denote hydrophobic sequences in the protein.

Table III

List of alkB-phoA and alkB-lacZ gene fusions and enzyme activity of the fusion proteins

Aminoacid Position[a]	Fusion Protein[b]	PhoA activity[c]	Fusion Protein[b]	LacZ activity[c]	Fusion point Location[d]
44	AP1	2	BG1	1760	C
52	11F	135	-	-	P
75	AP2	15	BG2	210	C
121	AP3	720	BG3	150	P
134	AP10	12	BG10	450	C
147	AP4	25	BG4	1630	C
184	AP5	19	BG5	880	C
229	AP11	35	BG11	190	C
251	AP6	420	BG6	45	P
265	2A	40	-	-	C
273	23G	25	-	-	C
287	AP7	3	BG7	810	C
311	8B	12	-	-	C
323	AP8	2	BG8	560	C
345	AP12	2	BG12	1560	C
389	AP9	1	BG9	1850	C

[a]Amino acid position of fusion joint in AlkB. [b] AP(x) and BG(x) protein fusions were constructed by site-directed mutagenesis. The remaining fusions were generated by random transposition of Tn*phoA*. [c]PhoA and LacZ activity are expressed in Miller units [d]'C' indicates a cytoplasmic location. 'P' indicates a periplasmic location.

shown in Table III. The alkaline phosphatase activities ranged from 3 to 720 Miller units. The ß-galactosidase activities ranged from 45 to 1850 Miller units. The activity of the AlkB-LacZ fusions appeared to correlate reasonably well with the amount of intact fusion protein found on Western blots, indicating that the level of activity depends on sensitivity to proteolysis.This is not the case for AlkB-PhoA fusion proteins. Inactive PhoA fusions are quite stable as judged by Western blotting (Fig. 2A).

Fig. 3 shows the enzyme activity of AlkB fusion proteins as a function of the fusion position. Fusion proteins with high alkaline phosphatase and low ß-galactosidase activities indicate enzyme fusions to periplasmic loops in AlkB (Michaelis *et al.*, 1983; Bassford *et al.* 1979; Froshauer *et al.*, 1988). Protein fusions at sites P3 and P6 identify two such loops, between hydrophobic segments 3 and 4, and 6 and 7, respectively.

Alkaline phosphatase fusions of low activity and corresponding active ß-galactosidase fusions are characteristic of cytoplasmic domains. Protein fusions at sites P1, P2, P10, P4, P5, P11, P7, P8, P12 and P9 identify such domains (Fig. 4).

All of the above results are mutually consistent with respect to alkaline phosphatase and ß-galactosidase activity for fusions at a specific position. The results are also consistent with respect to the properties of fusions within hydrophilic or hydrophobic regions, with the exception of fusions to site P1, which is located at the start of the second hydrophobic stretch (Fig. 1A, B).

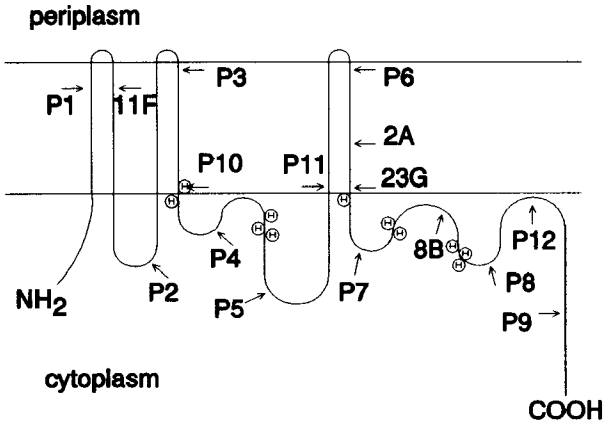

Fig. 4. **Proposed membrane topology of alkane hydroxylase.** The sites of Tn*phoA* insertion and *Pst*I sites introduced by site-directed mutagenesis to construct alkaline phosphatase and ß-galactosidase fusions are marked by arrows. The symbol H indicates histidines conserved between alkane hydroxylase and xylene monoxygenase.

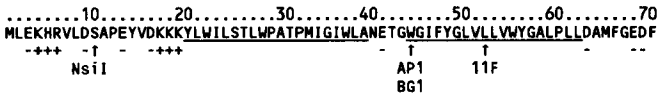

Fig. 5. **The amino terminal sequence of AlkB.** The two hydrophobic stretches are underlined. Charged residues are identified by a (+) or a (-). The location of the fusion joints of AP1, BG1 and 11F are indicated by the arrows.

Enzyme activity of fusions AP1, BG1 and 11F near the aminoterminus of AlkB- Alkaline phosphatase fusion 11F at residue 52 is quite active and it is processed like AP3 and AP6 (Fig. 2A), supporting a periplasmic localization of the region near residue 52. In contrast, the nearby PhoA fusion AP1 at residue 44 is completely inactive, despite the fact that considerable amounts of protein are produced (Fig. 2A, lane 2). The corresponding ß-galactosidase fusion BG1 is very active, and it is not degraded. Together, these data indicate a cytoplasmic location for position P1 and it appears therefore-that the hydrophobic region of AlkB which preceeds the P1 site is unable to translocate PhoA across the membrane. However, the additional hydrophobic AlkB sequence of amino acids 45 to 52 in the Tn*phoA* fusion 11F enable transport of PhoA across the cytoplasmic membrane. Position P1 must therefore be located more than halfway into the cytoplasmic membrane (Fig. 4).

The first hydrophobic stretch shows no export signal activity- The amino acid sequence of hydrophobic stretch 1 strongly resembles an export signal (Fig. 5). However, fusion AP1 is completely inactive and does not result in export of PhoA. To examine whether hydrophobic stretch 1 can act as an export signal or as a stop-transfer signal when it is in a different context, we moved it to different positions in the AlkB sequence. To this end we created an *Nsi*I site at amino acid position 10 in fusion AP1, and used

Table IV
List of AlkB-hydrophobic stretch 1-PhoA (AP(X)N) gene fusions and alkaline phosphatase activity of the fusion proteins

Protein fusion	PhoA Activity[a]	Protein fusion	PhoA Activity[a]
AP1	2		
AP3	380	AP3N	260
AP4	23	AP4N	25
AP6	190	AP6N	240
AP7	3	AP7N	4
AP8	2	AP8N	3
AP9	1	AP9N	2

[a]Alkaline phosphatase activity is expressed in Miller units.

this site to clone PhoA fused to hydrophobic stretch 1 (aminoacids 10-44) downstream of the hydrophobic segments 3, 4, 6, 7 and 8 using the P3, P4, P6, P7, P8 and P9 *PstI* sites. This results in a series of constructs which effectively contain hydrophobic stretch 1 inserted between the AlkB and PhoA portion of fusions AP3 to AP9 (Fig. 6). This allowed us to test the translocation activity of stretch 1, when it is preceded by an even or odd number of transmembrane stretches.

Table IV shows that the alkaline phosphatase activity of these constructs was only determined by the AlkB sequence preceeding hydrophobic stretch 1. The stretch itself had no effect on the localization of alkaline phosphatase, and it is incapable therefore of translocating the attached PhoA sequence through the cytoplasmic membrane, as illustrated for two constructs in Fig. 6.

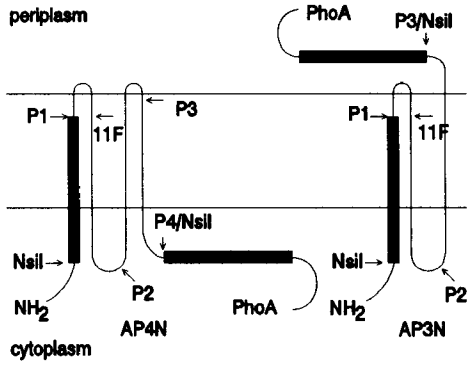

Fig. 6. **Structure of protein fusions AP4N and AP3N.** In PhoA fusions AP4N and AP3N, the first hydrophobic stretch, indicated by the filled bars, is inserted between the AlkB and PhoA domains of AP4 and AP3, using the *NsiI* and P1 *PstI* sites. The data in Table IV show that this stretch does not span the cytoplasmic membrane.

DISCUSSION

Alkane hydroxylase is an integral cytoplasmic membrane protein which contains eight hydrophobic sequences of sufficient length to be membrane-spanning segments (Kok *et al.*, 1989a). Since it is not possible to predict the topology of AlkB solely on the basis of its primary structure, we have prepared and assayed PhoA and LacZ fusions to AlkB in order to construct a model of the AlkB membrane topology.

Previous studies have shown that the fusion protein approach yields results which are consistent with other measures of membrane protein topology (Manoil and Beckwith, 1986). PhoA fusions to periplasmic domains exhibit high alkaline phosphatase activity (Hoffman and Wright, 1985; Manoil and Beckwith, 1985, 1986), while fusions to cytoplasmic domains generally show 10-100 fold-lower activities. The PhoA fusion approach has been successfully used for proteins with a single membrane spanning segment as well as for proteins spanning the membrane several times (reviewed in Manoil *et al.*, 1990). LacZ fusions generally show low activities if the fusion joint is facing the periplasm (Froshauer *et al.*, 1988) and consequently furnish complementary information.

Properties of AlkB-PhoA and AlkB-LacZ fusion proteins- Insertion of transposon Tn*phoA* into the target gene can be used to generate random alkaline phosphatase fusions. In our experiments, this approach alone did not allow a complete analysis of AlkB membrane topology. This was due to the occurrence of hot-spots, active out-of-frame fusions and the isolation of large numbers of active PhoA fusions outside AlkB. The method also has an inherent limitation in detecting fusions to cytoplasmic domains. Therefore, we decided to generate additional PhoA fusions using site-directed mutagenesis. This approach also allowed us to construct the corresponding LacZ fusions.

We found that both PhoA and LacZ fusions independently lead to the same model (Fig. 4), that is, the directed LacZ fusions BG3 and BG6 were the two least active fusions.

PhoA fusions to cytoplasmic domains varied considerably in activity. Fusions AP7, AP8 and AP9 to the carboxy terminal domain were completely inactive. This indicates that all information necessary for the cytoplasmic localization of the carboxy terminal domain is present in these fusions. In contrast, fusions AP2, AP10, AP4, AP5, AP11, 2A, and 23G, located near or in between transmembrane stretches, showed activities which were about 10-40 fold-lower than the activity of neighbouring periplasmic fusions, but were still clearly measurable. Apparently, these fusion proteins did not contain all of the necessary information for unequivocal cytoplasmic localization of the fusion joint. It has been shown that the alkaline phosphatase activity of PhoA fusions to the beginning of a hydrophilic cytoplasmic stretch can be much higher than fusions to the end of such stretches (Boyd *et al.*, 1987). This was attributed to the fact that such fusions lack subsequent localization determinants which are present in complete hydrophilic cytoplasmic loops, such as positively charged residues. These residues have been shown to strongly influence the translocation activity and orientation of a hydrophobic segment in the membrane (San Millan *et al.*, 1989; Boyd and Beckwith, 1989; Yamane *et al.*, 1990).

Protein fusions to hydrophobic stretch 1 and 2- The results obtained with the protein fusion approach were unambiguous for the region following hydrophobic stretch 2. This was not the case for the first two hydrophobic stretches. Based on the hydrophobicity plot (Fig. 1E), the presence of three lysines at the start of the first hydrophobic stretch and a negative charge at residue 41 (Fig. 1C, 5), we expected a periplasmic loop between the first and second membrane spanning segments near residues 40-42. However, the properties of the PhoA and LacZ fusions at residue 44 (P1) showed a cytoplasmic localization for PhoA and LacZ at this position. Eight additional amino acids resulted in a periplasmic localization of PhoA in fusion 11F at residue 52.

We tested the export signal and stop-transfer activity of the AlkB sequence from amino acids 10-44 (Fig. 5, 6) in different contexts, and found it to be negligible; this sequence can apparently neither effect nor block translocation of PhoA through the cytoplasmic membrane.

Thus, the first AlkB loop which is exposed to the periplasm is likely to be located closer to residue 52 than to residue 41.

Topology of AlkB- The model of Fig. 4 summarizes the above data. AlkB contains six transmembrane α-helices, exposing three short loops to the periplasm. Hydrophobic stretches 5 and 8 do not span the membrane. These stretches are located between fusions which have relatively low PhoA and high LacZ activities. With only 15 hydrophobic amino acids, bordered by charged residues, the hydrophobic stretch 5 is rather short, while stretch 8 differs from the other hydrophobic segments in its high proline content near the beginning of the stretch. This could limit the translocation activity of such a segment, perhaps because prolines interfere with the formation of an α-helical structure. In addition both stretches contain turn-promoting seg-

ments (Fig. 1D) according to the algorithm developed by Paul and Rosenbusch (1985), unlike segments 1, 2, 3, 4, 6 and 7.

Each of the six transmembrane stretches is more than twenty amino acids long. The last four transmembrane stretches contain no charged and few polar residues. The first transmembrane stretch must include the glutamic acid residue at position 41, and the second stretch must include the aspartic acid residue at position 63. A short amino terminal peptide (amino acids 1 to 20), two hydrophilic loops (amino acids 70 to 87 and 138 to 226) and a large carboxy terminal domain (amino acids 271 to 401) are positioned in the cytoplasm.

Enzyme activity of alkane hydroxylase- It is significant that the most distal AlkB-LacZ fusion (BG9) is able to restore growth on alkanes in an AlkB deletion strain (Table III). This shows that the active ß-galactosidase moiety does not interfere with the correct insertion of alkane hydroxylase into the membrane or with its interaction with rubredoxin, and that AlkB is in a (near) native conformation even though a bulky protein is attached to it.

At the moment it is not yet possible to speculate on the role of any of the hydrophilic segments in catalysis or in the interactions of AlkB with the other components of the alkane hydroxylase system. The AlkB protein is one of the few monooxygenases that are embedded in the cytoplasmic membrane by multiple hydrophobic stretches. In view of the substrate specificity of AlkB for n-alkanes it is possible that these water insoluble substrates reach the enzyme through the membrane. This suggests a general role for the hydrophobic segments in keeping the catalytic site in close proximity to the membrane.

Interestingly, the amino acid sequences of the xylene oxidase (XylM) of *Pseudomonas putida* mt-2 (TOL), which catalyzes the hydroxylation of methyl substituted aromatic compounds, and AlkB share 25% sequence identity (Suzuki *et al.*, 1991). It is not yet known whether xylene oxidase contains a ferrous iron as cofactor as is the case for alkane hydroxylase. However, it is striking that the sequence His-X_3-His occurs four times in both enzymes (Fig. 1F), as this sequence has been shown to chelate divalent metal ions, providing that they occur in an α-helix (Arnold and Haymore, 1991). A systematic isolation of mutations affecting the AlkB and XylM function, especially affecting substrate binding and specificity, should help to define the structural and functional properties of this class of monooxygenases.

ACKNOWLEDGEMENTS

We would like to thank Peter Terpstra for assisting with computer analysis of amino acid sequences, Jan Tommassen for providing anti-PhoA IgG, Greetje Vos-Scheperkeuter for providing anti-LacZ serum and Colin Manoil for providing Tn*phoA*, pCH2, pCH40, and the *E. coli* strains CC118 and LE372.